LAL 00-77
ROME1-1307/00
RM3-TH/00-16

2000 CKM-TRIANGLE ANALYSIS

A Critical Review with Updated Experimental Inputs and Theoretical Parameters

M. Ciuchini[(a)], G. D'Agostini[(b)], E. Franco[(b)], V. Lubicz[(a)], G. Martinelli[(b)], F. Parodi[(c)], P. Roudeau[(d)] and A. Stocchi[(d)]

[(a)] **Università di Roma Tre and INFN, Sezione di Roma III,**
Via della Vasca Navale 84, I-00146 Roma, Italy
[(b)] **Università "La Sapienza" and Sezione INFN di Roma,**
Piazzale A. Moro 2, 00185 Roma, Italy
[(c)] **Dipartimento di Fisica, Università di Genova and INFN**
Via Dodecaneso 33, 16146 Genova, Italy
[(d)] **Laboratoire de l'Accélérateur Linéaire**
IN2P3-CNRS et Université de Paris-Sud, BP 34, F-91898 Orsay Cedex

Abstract

Within the Standard Model, a review of the current determination of the sides and angles of the CKM unitarity triangle is presented, using experimental constraints from the measurements of $|\varepsilon_K|$, $|V_{ub}/V_{cb}|$, Δm_d and from the limit on Δm_s, available in September 2000. Results from the experimental search for $B^0_s - \bar{B}^0_s$ oscillations are introduced in the present analysis using the likelihood. Special attention is devoted to the determination of the theoretical uncertainties. The purpose of the analysis is to infer regions where the parameters of interest lie with given probabilities. The BaBar "*95% C.L. scanning*" method is also commented.

To be submitted to JHEP

1 Introduction

In the Standard Model, weak interactions of quarks are governed by the four parameters of the CKM matrix [1] which, in the Wolfenstein parametrisation [2], are labelled as λ, A, $\bar{\rho}$ and $\bar{\eta}$ [1]. Measurements of semileptonic decays of strange and beauty particles are the main sources of information on λ and A, respectively. The values of $|\varepsilon_K|$, $|V_{ub}/V_{cb}|$, Δm_d and Δm_s provide a set of four constraints for $\bar{\rho}$ and $\bar{\eta}$. These constraints depend, in addition, on other quantities obtained from measurements and/or theoretical calculations. The regions of $\bar{\rho}$ and $\bar{\eta}$ preferred by the four constraints are expected to overlap, as long as the Standard Model gives an overall description of the various experimental observations.

Since several years there has been intense activity to constrain the allowed region in the $(\bar{\rho}, \bar{\eta})$ plane from the best knowledge of the experimental and theoretical inputs [3]–[18]. Though the analysis methods differ in some details, they have a common ground in what we shall call *standard approach* through this paper. First, the goal of the various authors has been, explicitly or implicitly, to infer regions in which the values of $\bar{\rho}$ and $\bar{\eta}$ are contained with a certain level of probability (or *confidence*). Second, uncertainties due to statistical errors and systematic effects in experiments, as well as theoretical uncertainties, are combined together to deduce a global uncertainty about $\bar{\rho}$ and $\bar{\eta}$. As far as this second point is concerned, the various authors have used different "prescriptions" which can be seen, indeed, as approximations of the consistent Bayesian method which is described, and adopted, in this paper. The 68% probability regions favoured by the data and the theoretical understanding of the relevant processes select quite narrow regions for $\bar{\rho}$ and $\bar{\eta}$, largely independent of the details of the specific methods and of the different treatment of (experimental) systematic and theoretical uncertainties.

A different approach, named "*95% C.L. scanning*" in this paper[2], has been adopted in the BaBar Physics Book [21], and recently used in [22]. In this approach, it is stated that it is not possible to define probability distributions for theoretical parameters coming from calculations affected by systematic uncertainties or based on educated guesses (in practice all theoretical parameters and some experimental systematics belong to this class). On the basis of these considerations, the *95% C.L. scanning* approach rejects the two basic points of the standard method and a different procedure is proposed. For the theoretical inputs, it is assumed that one can only define intervals inside which the *true* values of the parameters are contained. At fixed values of the theoretical inputs (within the allowed intervals) a maximum likelihood fit, which includes the other sources of uncertainties, is made, and 95% C.L. contours are determined. Finally, the envelope of such contours is "*proposed to be a (conservative) method to obtain some 95% C.L. regions for all CKM parameters*" [22]. Using the same arguments of the BaBar Physics Book [21], this procedure has been recently recommended in [23, 24], as opposed to the standard approach which is claimed of being too optimistic.

In view of the importance of constraining the parameters of the CKM matrix, or of the possibility of detecting signals of new physics in low-energy weak decays, in this paper we reconsider the whole matter, and in particular we focus on the most critical issues of the CKM-triangle analysis, namely the uncertainty on theoretical parameters and the inferential framework to handle consistently all uncertainties. This also allows to answer

[1] $\bar{\rho} = \rho(1 - \frac{\lambda^2}{2})$ and $\bar{\eta} = \eta(1 - \frac{\lambda^2}{2})$ [3].

[2] Note that this *95% C.L. scanning* is different from the "scanning method" used to predict e.g. ε'/ε [19] (see also [20] for comments).

to several important controversial questions raised in the past, namely:

- whether it is possible, or necessary, to assign a probability distribution function (p.d.f.) to theoretical parameters;
- whether it is possible to define a p.d.f. for quantities extracted from physical measurements and from theoretical parameters affected by systematic uncertainties;
- whether average values, errors and p.d.f. of the quantities considered in the present analysis depend in a crucial way on the assumptions made on the theoretical parameters and the corresponding errors.

The main conclusions of this study are the following:

1. The standard method is a theoretically sound approach, which finds its justification within the inferential framework discussed in this paper. This method allows a consistent treatment of the systematic and theoretical uncertainties and makes it possible to define regions where the values of $\bar{\rho}$ and $\bar{\eta}$ (as well as of other quantities of phenomenological interest) are contained with any given level of confidence. In this respect the criticisms of [23, 24] are not justified.

2. The BaBar *95% C.L. scanning*, instead, is based on an *ad hoc* prescription intended to define *only* a "95% C.L." region. The meaning of this statement is unclear: the so called "95% C.L. region" does not correspond to the usual statistical definition of 95% confidence that the values of parameters lie in that region, neither in a frequentist sense, nor in a Bayesian one.

3. For the sake of comparison, with the standard method we have used the same values of the input parameters and tried to mimic the same uncertainties of the *95% C.L. scanning*. We find that the "95% C.L." regions selected in the ($\bar{\rho}$, $\bar{\eta}$) plane with the two methods are very similar, and it is thus not founded to qualify as too optimistic the standard approach.

4. In the *95% C.L. scanning* approach the information contained in the p.d.f. of the relevant quantities ($\bar{\rho}$, $\bar{\eta}$, $\sin(2\beta)$, etc.) is missing. Thus we only know that a certain quantity is somehow expected in a given interval, but we do not know which is the most probable value, what is the shape of its p.d.f. etc. We show, instead, that this important information can be extracted from the data, using the standard method, in spite of the uncertainties in the theoretical parameters.

Regarding the analysis, since this kind of studies have been extensively illustrated in previous publications, see for example [13] and [17], here we only discuss in detail $B_s^0-\bar{B}_s^0$ mixing, for which we adopted a different procedure, based directly on the likelihood (Section 6).

The main results for the physical quantities ($\bar{\rho}$, $\bar{\eta}$, $\sin(2\beta)$, etc.) can be found in Sections 7.

The remainder of the paper is organised as follows. In Section 2 we summarise the theoretical constraints between $\bar{\rho}$ and $\bar{\eta}$ and the available experimental and theoretical inputs in the Standard Model. In Section 3 we describe the inferential framework used in this study and relate it to different standard approach analyses. Our comments on the *95% C.L. scanning* method are also presented in this Section. The choice of values and uncertainties for the most critical theoretical parameters is discussed in Section 4. In Section 5 the p.d.f. determination for $|V_{cb}|$, $|V_{ub}|$ and the parameter λ are explained. A new method to include the information coming from searches of $B_s^0 - \bar{B}_s^0$ mixing is illustrated in Section 6. The results of the analysis are presented and discussed in Section 7. The stability of the results has been verified by varying the different input parameters in Section 8. A comparison of our results with those obtained with *95% C.L. scanning* is made in Section 9. Finally, conclusions are drawn in Section 10.

2 Standard Model formulae relating $\bar{\rho}$ and $\bar{\eta}$ to experimental and theoretical inputs

Four measurements restrict, at present, the possible range of variations of the $\bar{\rho}$ and $\bar{\eta}$ parameters:

- The relative rate of charmed and charmless b-hadron semileptonic decays which allows to measure the ratio

$$\left|\frac{V_{ub}}{V_{cb}}\right| = \frac{\lambda}{1 \; - \; \frac{\lambda^2}{2}} \sqrt{\bar{\rho}^2 + \bar{\eta}^2}\,. \tag{1}$$

.

- The $B_d^0 - \bar{B}_d^0$ time oscillation period which can be related to the mass difference between the light and heavy mass eigenstates of the $B_d^0 - \bar{B}_d^0$ system

$$\Delta m_d \;=\; \frac{G_F^2}{6\pi^2} m_W^2 \; \eta_c S(x_t) \; A^2\lambda^6 \; [(1-\bar{\rho})^2 + \bar{\eta}^2] \; m_{B_d} \; f_{B_d}^2 \hat{B}_{B_d} \;, \tag{2}$$

where $S(x_t)$ is the Inami-Lim function [25] and $x_t = m_t^2/M_W^2$. m_t is the $\overline{MS}$ top mass, $m_t^{\overline{MS}}(m_t^{\overline{MS}})$, and η_c is the perturbative QCD short-distance NLO correction. The remaining factor, $f_{B_d}^2 \hat{B}_{B_d}$, encodes the information of non-perturbative QCD. Apart for $\bar{\rho}$ and $\bar{\eta}$, the most uncertain parameter in this expression is $f_{B_d}\sqrt{\hat{B}_{B_d}}$. The value of $\eta_c = 0.55 \pm 0.01$ has been obtained in [26] and we used $m_t = (167 \pm 5)$ GeV, as deduced from measurements of the mass by CDF and D0 Collaborations [27].

- The limit on the lower value for the time oscillation period of the $B_s^0 - \bar{B}_s^0$ system is transformed into a limit on Δm_s and compared with Δm_d

$$\frac{\Delta m_d}{\Delta m_s} \;=\; \frac{m_{B_d} f_{B_d}^2 \hat{B}_{B_d}}{m_{B_s} f_{B_s}^2 \hat{B}_{B_s}} \; \lambda^2 \; [(1-\bar{\rho})^2 + \bar{\eta}^2]\,. \tag{3}$$

The ratio $\xi \;=\; f_{B_s}\sqrt{\hat{B}_{B_s}}/f_{B_d}\sqrt{\hat{B}_{B_d}}$ is expected to be better determined from theory than the individual quantities entering into its expression. In our analysis, we accounted for the correlation due to the appearance of Δm_d in both Equations (2) and (3).

- CP violation in the kaon system which is expressed by $|\varepsilon_K|$

$$|\varepsilon_K| \ = C_\varepsilon \ A^2\lambda^6 \ \bar{\eta} \left[-\eta_1 S(x_c) + \eta_2 S(x_t) \left(A^2\lambda^4 \left(1-\bar{\rho}\right)\right) + \eta_3 S(x_c, x_t)\right] \ \hat{B}_K \ , \quad (4)$$

where

$$C_\varepsilon = \frac{G_F^2 f_K^2 m_K m_W^2}{6\sqrt{2}\pi^2 \Delta m_K} \ . \quad (5)$$

$S(x_i)$ and $S(x_i, x_j)$ are the appropriate Inami-Lim functions [25] of $x_q = m_q^2/m_W^2$, including the next-to-leading order QCD corrections [26, 28]. The most uncertain parameter is $\hat{B}_K$.

Constraints are obtained by comparing present measurements with theoretical expectations using the expressions given above and taking into account the different sources of uncertainties. In addition to $\bar{\rho}$ and $\bar{\eta}$, these expressions depend on other quantities which have been listed in Table 2. Additional measurements or theoretical determinations have been used to provide information on the values of these parameters.

3 Inferential framework

In this Section we recall the basic ingredients of the standard method, interpreted in the framework of the Bayesian approach. This allows us to discuss the role of the systematic and theoretical uncertainties in deriving probability intervals for the relevant parameters.

3.1 Standard approach and Bayesian inference

Each of Equations (1)–(4) relates a constraint c_j (where c_j stands for $|V_{ub}/V_{cb}|$, Δm_d, $\Delta m_d/\Delta m_s$ and $|\varepsilon_K|$, for $j = 1, \ldots, 4$) to the CKM–triangle parameters $\bar{\rho}$ and $\bar{\eta}$, via the set of ancillary parameters $\mathbf{x}$, where $\mathbf{x} = \{x_1, x_2, \ldots, x_N\}$ stand for all experimentally determined or theoretically calculated quantities from which the various c_j depend

$$c_j = c_j(\bar{\rho}, \bar{\eta}; \mathbf{x}). \quad (6)$$

In an ideal case of exact knowledge of c_j and $\mathbf{x}$, each of the constraints provides a curve in the $(\bar{\rho}, \bar{\eta})$ plane. In such a case, there would be no reason to favour any of the points on the curve, unless we have some further information or physical prejudice, which might exclude points outside a determined *physical region*, or, in general, assign different weights to different points. In a realistic case, we suffer from several uncertainties on the quantities c_j and $\mathbf{x}$. Uncertainty does not imply, however, that we are absolutely ignorant about a given quantity. First of all, there are values which, to the best of our knowledge, we consider ruled out (for example a value of m_t of 100 GeV or 500 GeV). Second, we assign different probabilities to the values within the "almost certain range", $147\,\text{GeV} < m_t < 187\,\text{GeV}$ say [3]. In the m_t case, for example, we think that it is much more probable that the value of m_t lies between 157 and 177 GeV rather than in the rest of the interval, in spite of the fact that the two sub-intervals have the same widths.

This means that, instead of a single curve (6) in the $(\bar{\rho}, \bar{\eta})$ plane, we have a family of curves which depends on the distribution of the set $\{c_j, \mathbf{x}\}$. As a result, the points in the

[3] In this example m_t is the $\overline{MS}$ top mass of Equation (2), $m_t = (167 \pm 5)$ GeV.

Parameter	Value	Gaussian σ	Uniform half-width	Ref.
λ	0.2237	0.0033, see text		sect. 5.3 (eq. 46)
$\|V_{cb}\|$	41.0×10^{-3}	1.6×10^{-3}, see text		sect. 5.1 (eq. 43)
$\|V_{ub}\|$	35.5×10^{-4}	3.6×10^{-4}	–	sect. 5.2 (eq. 45)
$\|\varepsilon_K\|$	2.280×10^{-3}	0.019×10^{-3}	–	[29]
Δm_d	0.487 ps^{-1}	0.014 ps^{-1}	–	[30]
Δm_s	$>$ 15.0 ps^{-1} at 95% C.L.	see text		[30]
m_t	167 GeV	5 GeV	–	[27]
m_b	4.23 GeV	0.07 GeV	–	[31]
m_c	1.3 GeV	0.1 GeV	–	[29]
$\hat{B}_K$	0.87	0.06	0.13	sect. 4 (eq. 34)
$f_{B_d}\sqrt{\hat{B}_{B_d}}$	230 MeV	25 MeV	20 MeV	sect. 4 (eq. 28)
$\xi = \frac{f_{B_s}\sqrt{\hat{B}_{B_s}}}{f_{B_d}\sqrt{\hat{B}_{B_d}}}$	1.14	0.04	0.05	sect. 4 (eq. 31)
α_s	0.119	0.03	–	[28]
η_1	1.38	0.53	–	[28]
η_2	0.574	0.004	–	[28]
η_3	0.47	0.04	–	[28]
η_c	0.55	0.01	–	[28]
f_K	0.161 GeV	fixed		[29]
Δm_K	0.5301×10^{-2} ps^{-1}	fixed		[29]
G_F	1.16639×10^{-5}GeV^{-2}	fixed		[29]
m_W	80.41 GeV	fixed		[29]
$m_{B_d^0}$	5.2792 GeV	fixed		[29]
$m_{B_s^0}$	5.3693 GeV	fixed		[29]
m_K	0.493677 GeV	fixed		[29]

Table 1: Values of the quantities entering into the expressions of $|\varepsilon_K|$, $|V_{ub}/V_{cb}|$, Δm_d and Δm_s. In the third and fourth columns the Gaussian and the flat part of the uncertainty are given, respectively.

$(\bar{\rho}, \bar{\eta})$ plane get different weights (even if they were taken to be equally probable *a priori*) and our *confidence* on the values of $\bar{\rho}$ and $\bar{\eta}$ clusters in a region of the plane.

The above arguments, which we consider very natural and close to physicist intuition, can be formalised by using the so called Bayesian approach (see [32] for an introduction). In this approach, the uncertainty is described in terms of a probability density function $f(\cdot)$, which quantifies our confidence on the values of a given quantity. The inference of $\bar{\rho}$ and $\bar{\eta}$ becomes then a straightforward application of probability theory, getting rid of all "*ad hoc* prescriptions".

The simplest way to implement the probabilistic reasoning discussed above is to define an idealised "p.d.f." for each constraint

$$f(\bar{\rho}, \bar{\eta} \,|\, c_j, \mathbf{x}) \propto \delta(c_j - c_j(\bar{\rho}, \bar{\eta}, \mathbf{x}))\,, \tag{7}$$

where δ is the Dirac delta distribution. The quotes recall us that this p.d.f. is a distribution in a mathematical sense, which is to be taken as the limit of a very narrow p.d.f.

with values different from zero only along a curve. The p.d.f. which takes into account the full uncertainty about c_j and $\mathbf{x}$ is obtained from (7) by making use of the standard probability rules

$$\begin{aligned}
f(\bar{\rho},\bar{\eta}) &= \int f(\bar{\rho},\bar{\eta}\,|\,c_j,\mathbf{x})\cdot f(c_j,\mathbf{x})\,\mathrm{d}c_j\,\mathrm{d}\mathbf{x} && (8)\\
&\propto \int \delta(c_j - c_j(\bar{\rho},\bar{\eta},\mathbf{x}))\cdot f(c_j)\cdot f(\mathbf{x})\,\mathrm{d}c_j\,\mathrm{d}\mathbf{x} && (9)\\
&\propto \int \delta(c_j - c_j(\bar{\rho},\bar{\eta},\mathbf{x}))\cdot \frac{1}{\sqrt{2\pi}\,\sigma(c_j)}\exp\left[-\frac{(c_j-\hat{c}_j)^2}{2\,\sigma^2(c_j)}\right]\cdot f(\mathbf{x})\,\mathrm{d}c_j\,\mathrm{d}\mathbf{x} && (10)\\
&\propto \int \frac{1}{\sqrt{2\pi}\,\sigma(c_j)}\exp\left[-\frac{(c_j(\bar{\rho},\bar{\eta},\mathbf{x})-\hat{c}_j)^2}{2\,\sigma^2(c_j)}\right] f(x_1)\cdot f(x_2)\cdots f(x_N)\,\mathrm{d}\mathbf{x}\,, && (11)
\end{aligned}$$

where $\hat{c}_j$ is the experimental best estimate of c_j, with uncertainty $\sigma(c_j)$. A Gaussian distribution has been assumed just for simplicity and without lack of generality. The joint p.d.f. $f(c_j,\mathbf{x})$ has been splitted as a product of the individual p.d.f., assuming the independence of the different quantities, which is a very good approximation for the case under study.

As alternative procedure, one may introduce a global inference relating $\bar{\rho}$, $\bar{\eta}$, c_j and $\mathbf{x}$. This is followed by a second step where marginalization is performed over those quantities which we are not interested to. In this case, by making use of Bayes' theorem, we obtain

$$\begin{aligned}
f(\bar{\rho},\bar{\eta},c_j,\mathbf{x}\,|\,\hat{c}_j) &\propto f(\hat{c}_j\,|\,c_j,\bar{\rho},\bar{\eta},\mathbf{x})\cdot f(c_j,\bar{\rho},\bar{\eta},\mathbf{x}) && (12)\\
&\propto f(\hat{c}_j\,|\,c_j)\cdot f(c_j\,|\,\bar{\rho},\bar{\eta},\mathbf{x})\cdot f(\mathbf{x},\bar{\rho},\bar{\eta}) && (13)\\
&\propto f(\hat{c}_j\,|\,c_j)\cdot \delta(c_j - c_j(\bar{\rho},\bar{\eta},\mathbf{x}))\cdot f(\mathbf{x})\cdot f_\circ(\bar{\rho},\bar{\eta})\,, && (14)
\end{aligned}$$

where $f_\circ(\bar{\rho},\bar{\eta})$ denotes the *prior* distribution as often used in the literature. The various steps follow from probability rules, by assuming the independence of the different quantities and by noting that $\hat{c}_j$ depends on $(\bar{\rho},\bar{\eta},\mathbf{x})$ only via c_j. This is true since c_j is univocally determined, within the Standard Model, from the values of $\bar{\rho}$, $\bar{\eta}$ and $\mathbf{x}$ (hence the limit to a delta function of its p.d.f.). We then recover Equation (11): i) by assuming a Gaussian error function for $\hat{c}_j$ around c_j; ii) by considering the various x_i as independent; iii) by taking a flat a priori distribution for $\bar{\rho}$ and $\bar{\eta}$ and iv) by integrating Equation (14) over c_j and $\mathbf{x}$. Note that equiprobability of all points in the $(\bar{\rho},\bar{\eta})$ plane was also implicit in Equation (9), as discussed above.

Although the first derivation of Equation (9) is probably the most intuitive one, hereafter we use the second one, which is the usual way of performing Bayesian inference. The second procedure also shows explicitly the connection with the methods which we denoted as standard in the introduction.

The extension of the formalism to several constraints is straightforward. We can rewrite Equation (12) as

$$f(\bar{\rho},\bar{\eta},\mathbf{x}\,|\,\hat{c}_1,...,\hat{c}_{\mathrm{M}}) \;\propto\; \prod_{j=1,\mathrm{M}} f_j(\hat{c}_j\,|\,\bar{\rho},\bar{\eta},\mathbf{x}) \times \prod_{i=1,\mathrm{N}} f_i(x_i)\times f_\circ(\bar{\rho},\bar{\eta})\,. \qquad (15)$$

In the derivation of (15), we have used the independence of the different quantities. Moreover, the conditioning from $f_j(\cdot)$ on the c_j have been removed, since the c_j act as intermediate variables which are finally integrated away. The derivation (8)–(11) can also be easily extended to the case of several constraints and leads, again, to the same result as that found by using Bayes' theorem. We have only to account properly the weight on c_j, induced by the other constraint(s) previously considered. In this case Equation (7) becomes $f(\bar{\rho}, \bar{\eta} \,|\, c_1, \ldots, c_M, \mathbf{x}) \propto \delta(c_M - c_M(\rho, \eta, \mathbf{x})) \cdot f(\rho, \eta \,|\, c_1, \ldots, c_{M-1}, \mathbf{x})$.

By integrating Equation (15) over $\mathbf{x}$ we can rewrite the inferential scheme in the following convenient way

$$f(\bar{\rho}, \bar{\eta} \,|\, \hat{\mathbf{c}}, \mathbf{f}) \propto \mathcal{L}(\hat{\mathbf{c}} \,|\, \bar{\rho}, \bar{\eta}, \mathbf{f}) \times f_\circ(\bar{\rho}, \bar{\eta}) \,, \tag{16}$$

where $\hat{\mathbf{c}}$ stands for the set of measured constraints, and

$$\mathcal{L}(\hat{\mathbf{c}} \,|\, \bar{\rho}, \bar{\eta}, \mathbf{f}) = \int \prod_{j=1,\mathrm{M}} f_j(\hat{c}_j \,|\, \bar{\rho}, \bar{\eta}, \mathbf{x}) \prod_{i=1,\mathrm{N}} f_i(x_i) \, \mathrm{d}x_i \tag{17}$$

is the effective overall likelihood which takes into account all possible values of x_j, properly weighted. We have written explicitly that the overall likelihood depends on the best knowledge of all x_i, described by $f(\mathbf{x})$.

Whereas *a priori* all values for $\bar{\rho}$ and $\bar{\eta}$ are considered equally likely, *a posteriori* the probability clusters around the point which maximises the likelihood. This is the reason why, in principle, different procedures for determining $\bar{\rho}$ and $\bar{\eta}$, based on the maximum likelihood, are equivalent to the method described here and should get similar results. We say "in principle" because other methods are typically implemented using the χ^2 minimisation. This implies the assumption of a multi-Gaussian solution of the integral (17), with overall standard deviations which are simply a quadratic combinations of the "uncertainties" related to each x_i. On the other hand, a quadratic combination relies on the approximative linear dependence of c_j from the possible variations of x_i.

In conclusion, the final (unnormalised) p.d.f. obtained starting from a flat distribution of $\bar{\rho}$ and $\bar{\eta}$ is

$$f(\bar{\rho}, \bar{\eta}) \propto \int \prod_{j=1,\mathrm{M}} f_j(\hat{c}_j \,|\, \bar{\rho}, \bar{\eta}, \mathbf{x}) \prod_{i=1,\mathrm{N}} f_i(x_i) \, \mathrm{d}x_i \,. \tag{18}$$

The integration can be done by Monte Carlo methods, the normalisation is trivial, and all moments can be calculated in a (conceptually) easy way. Obviously there are several ways to implement the Monte Carlo integration, using different techniques to generate events. A comparison of the results obtained with the approach of [17], where some effort has been done to improve the generation efficiency, and the results of [13] is presented in Section 7.

It is important to note that the inferential method does not make any distinction on whether the individual likelihood associated to some constraint is different from zero only in a narrow region (and we usually refer to this case as "measurement"), or it goes to zero only on one of the two sides (e.g. when $c_j \to \infty$ or 0). In the latter case, the data only provide an upper/lower bound to the value of the constraint. This is precisely what happens, at present, with Δm_s. Therefore, the experimental information about this constraint enters naturally in the analysis (more details can be found in Section 6).

3.2 Treatment of systematic and theoretical uncertainties

At this point, it is in order a discussion on some important ingredients of the analysis which raised some controversy in the past. They are related to the quantitative handling of the uncertainties due to systematic effects and to theoretical inputs.

In Equation (9) we have written explicitly that $\hat{c}_j$ is Gaussian distributed around c_j. As a consequence, we tend to say that also c_j is Gaussian distributed around $\hat{c}_j$, the inversion being well understood in the case of random errors. The question is how to include the case where also systematic uncertainties are present. One of the nice features of the Bayesian approach is that the uncertainty has positively the same meaning, and there is no conceptual distinction between the uncertainty due to random fluctuations, which might have occurred in the measuring process, the uncertainty about the parameters of the theory, and the uncertainty about *influence quantities* (i.e. "systematics") of not-exactly-known value, (see [32] and [33]). Under the assumption that the individual likelihoods $f_j(\cdot)$ of Equations (15)–(18) do only depend on random effects, the uncertainty due to systematics can be included using, again, concepts and formulae of conditional probability. In fact, calling $\mathbf{h}$ the set of influence quantities on which the measured constraints may depend, with joint p.d.f. $g(\mathbf{h})$, the likelihood (17) becomes

$$\mathcal{L}(\hat{\mathbf{c}}\,|\,\bar{\rho},\bar{\eta},\mathbf{f},\mathbf{g}) = \int \prod_{j=1,\mathrm{M}} f_j(\hat{c}_j\,|\,\bar{\rho},\bar{\eta},\mathbf{x},\mathbf{h}) \cdot f(\mathbf{x}) \cdot g(\mathbf{h})\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{h}\,, \tag{19}$$

where we have written the p.d.f. of $\mathbf{x}$ in its general form, allowing also correlations among the elements. As can be seen from Equation (19), there is neither a conceptual nor a formal distinction between the handling of $\mathbf{x}$ and $\mathbf{h}$. Therefore, we can simply extend the notation to include in $\mathbf{x}$ the influence parameters responsible of the systematic uncertainty, and use Equation (17) in an extended way. This is what it has been actually done in the past to infer $f(\bar{\rho},\bar{\eta})$ with the Monte Carlo integration method resulting from (18). Moreover, see also [33] and references therein, we arrive to the following conclusion. Irrespectively of the assumptions made on the p.d.f. of $\mathbf{x}$, the overall likelihoods $f(\hat{c}_j)$ are approximately Gaussian because of a mechanism similar to the central limit theorem (i.e. just a matter of combinatorics). This makes the results largely stable against variations within *reasonable* choices of models and parameters used to describe the uncertainties due to theory and systematics. This also explains why methods based on χ^2 minimisation (for example refs. [15] and [16]) can be considered as approximations of the one used in refs. [12],[13], [14] and [17]. We stress again that a common ground of all the methods that we classify as standard is to produce regions where $\bar{\rho}$ and $\bar{\eta}$ are contained with any given level of confidence. On the contrary, the BaBar *95% C.L. scanning* is based on an *ad hoc* prescription which obscures the meaning of the results. In that approach, a so called "95% C.L." is produced, which does not correspond to the usual 95% confidence that the parameters lie in that regions.

As far as the choice of the mathematical expression for $f_i(x_i)$ is concerned, practical examples of simple models can be found in [33]. Due to the insensitivity of the result on the precise model, as discussed previously and as shown later in this paper, we simplify the problem, by reducing the choice only to two possibilities. We choose a Gaussian model when the uncertainty is dominated by statistical effects, or there are many comparable contributions to the systematics error, so that the central limit theorem applies. We choose a uniform p.d.f. if the parameter is believed to be (almost) certainly in a given interval, and the points inside this interval are considered equally probable.

A final comment, before ending this Section, concerns the compatibility among inferences provided by individual constraints. In the simplified approach based on χ^2 minimisation, a conventional evaluation of compatibility stems automatically from the value of the χ^2 at its minimum. This information is lost in the likelihood approach, but we accept this loss without any regret, in view of what we gain. It is well known, indeed, that crude arguments about compatibility or incompatibility, based only on the minimum value of the χ^2 and the number of degrees of freedom lead to misleading conclusions (as premature claims of new physics have demonstrated in the past decades). Therefore, we consider more reasonable to judge the compatibility of the constraints by comparing partial inferences obtained when removing each constraint at the time. Examples are given in Figures 8 and 10. In our analysis, the overlap of the various constraints is excellent, and therefore we have no reason to suspect deviations from the Standard Model, given the available experimental information.

4 Theoretical inputs: B_K, $f_B\sqrt{\hat{B}_B}$ and the ξ parameter

In this Section we discuss the theoretical inputs which have been used in our phenomenological analysis. In particular, we explain how central values and uncertainty models for the different quantities have been chosen. This gives us the opportunity of clarifying some issues on which there is, we think, some confusion in the literature. This discussion may be instructive especially for non-lattice experts (often experimentalists) who are engaged in this kind of analyses and have to find some orientation in using results from a *plethora* of lattice studies.

For example, in ref. [18] one finds statements like "*Some recent lattice computations of heavy meson decay constants are shown in Table IX... Moreover they have been computed in a renormalization scheme which is peculiar to the lattice gauge theory, and have to be converted using a continuum renormalization scheme (such as the $\overline{MS}$ scheme), within which the experimental data are analysed.*" This statement is obviously *wrong*: decay constants (which are related to matrix elements of the weak axial current), as all measurable physical quantities, are *scheme independent*. It is not possible to compute any matrix element of the axial current in the $\overline{MS}$ scheme simply because such a quantity does not exist. Indeed the *physical* scheme-independent axial current is obtained from the lattice one by a finite renormalization constant, $A_\mu = Z_A A_\mu^{latt}$, which can be (has been) determined non-pertubatively with a negligible uncertainty[34, 35].

Scheme dependence only enters some theoretical predictions (but never in those for the decay constants) because the perturbative calculations of the Wilson coefficients in the effective Hamiltonians, relevant for weak decays and mixing, are truncated at a certain order (typically NLO). This is not a peculiarity of the lattice approach but it depends on the limited number of orders which has been computed in continuum perturbation theory. Moreover, the choice of presenting the calculations in $\overline{MS}$ is only a traditional option which has not to do with the continuum or the lattice formulation of the theory. A discussion of this point in the case of B^0–$\bar{B}^0$ mixing can be found below.

Another source of confusion is the uncontrolled propagation of values and errors of the parameters from one review talk to another, without verification on the original papers, and regardless of more recent calculations and progresses. A typical example is the value $\xi \sim 1.3$ quoted in [36]. This number was only found in [37]. All other lattice calculations find for this quantity values between 1.11 and 1.17 [38]–[41] in the quenched approximation and similar numbers have been recently confirmed by unquenched data [41]. Similarly, in [23], A. Falk quotes $\xi = 1.14(0.13)$ on the basis of a two-year old review by S. Sharpe [42]. In the absence of unquenched calculations, the rather generous uncertainty reported in [42] was justified, at the time, on the basis of theoretical estimates obtained by using quenched and unquenched chiral perturbation theory. These estimates have not been confirmed by (partially) unquenched results, which were already available last year [4]. Irrespectively of recent lattice progresses, this very large uncertainty, which is not supported by any explicit numerical result, risks to survive and be used in future phenomenological analyses. The most recent figures for the B-meson decay constants, the B-parameters, $f_{B_d}\sqrt{\hat{B}_{B_d}}$ and ξ can be found in Tables 2 and 3. These tables include the results presented at Lattice 2000 [41]. In the following, for all the theoretical parameters, we have used results taken from lattice QCD. There are several reasons for this choice, which has been adopted also in previous studies of the unitarity triangle [4, 7, 12, 13, 17]. Lattice QCD is not a model, as the quark model for example, and therefore physical quantities can be computed from first principles without arbitrary assumptions. It provides a method for predicting all physical quantities (decay constants, weak amplitudes, form factors) within a unique, coherent theoretical framework. For many quantities the statistical errors have been reduced to the percent level (or even less). Although most of the results are affected by systematic effects, the latter can be "systematically" studied and eventually corrected. All the recent literature on lattice calculations is indeed focused on discussions of the systematic errors and studies intended to reduce these sources of uncertainty. We are not aware of any other approach (1/N expansion, QCD Sum Rules, etc.) where such a deep investigation of systematic errors is being carried out for a so large set of physical quantities as in lattice QCD. Finally, in cases where predictions (non post-dictions) from lattice QCD have been compared with experiments, for example f_{D_s}, the agreement has been found very good.

Obviously, for some quantities the uncertainty from lattice simulations is far from being satisfactory and further effort is needed to improve the situation. For the quantities considered here this is particularly true for the B^0_d–$\bar{B}^0_d$ mixing amplitude, as discussed below. Nevertheless, for the reasons mentioned before, we think that lattice results and uncertainties are the most reliable ones and we have used them in our study.

4.1 Statistical and systematic effects in lattice calculation uncertainties

Lattice simulations are theoretical experiments carried out by numerical integration of the functional integral by Monte Carlo techniques. In this respect uncertainties are evaluated following criteria very close to those used in experimental measurements. The results are obtained with "statistical errors", i.e. uncertainties originated by stochastic fluctuations, which may be reduced by increasing the sample of gluon-field configurations on which the averages are performed. It is very reasonable to assume that the statistical fluctuations have a Gaussian (almost Gaussian) distribution. Hence, the probabilistic inversion needed to infer the quantity of interest gives rise to Gaussian uncertainty models.

To convert the results of lattice simulations in predictions for the physical amplitudes several steps are necessary:

[4] We denote as *partially* unquenched results those obtained with two sea-quark flavours at values of the light-quark masses larger than the physical ones, typically of the order of the strange quark mass, or slightly below.

a) renormalization of the relevant operators;

b) extrapolation to the continuum limit, namely to zero lattice spacing ($a \to 0$);

c) unquenched calculations. The most precise numbers have been obtained in the quenched approximation. Theoretical estimates and some preliminary results in the (partially) unquenched case are also known and they are used to estimate the systematic errors of the quenched results.

a)–c) are the main sources of systematic errors for B_K, $f_{B_q}\sqrt{\hat{B}_{B_q}}$ and ξ and are discussed separately in the following. Here we want only to stress that systematic errors from lattice calculations are conceptually similar to some of the systematic errors present in experimental measurements.

Let us consider as an example discretization errors. Since it is obviously impossible to work at zero lattice spacing, a method to correct for discretization effects is to compute a given physical quantity at different values of the lattice spacing and to extrapolate it to zero lattice spacing. The theory tells us whether the extrapolation has to be linear or quadratic in a. Thus for example one fits a given quantity Q as

$$Q(a) = Q(a=0) + Q'a + Q''a^2 + \dots , \tag{20}$$

and takes $Q(a=0)$ as best estimate of the value of Q in the continuum limit. Since the "measurements" performed at fixed lattice spacing are subject to a statistical error, the uncertainty in the extrapolated quantity is inflated with respect to the points directly measured. In the quenched case, extrapolations have been made for both $\hat{B}_K$, the B-meson decay constants and ξ. Systematic studies for the B-meson mixing parameters are still missing. From a comparison among calculations performed by different groups at different lattice spacing and with different lattice actions, an estimate of discretization errors can be obtained, however, also in these cases.

When an extrapolation to the continuum limit of the lattice data has been possible, the final uncertainty results from the statistical error of the points measured at fixed lattice spacing (a residual uncertainty is present when linear or linear plus quadratic extrapolations give different results). Thus, in this case, it is natural to assume that the final error has a Gaussian distribution.

As for the errors coming from quenched calculations, partially unquenched calculations exist for several quantities considered in this study, namely $\hat{B}_K$ and $f_{B_{d,s}}$. These calculations are usually performed with two light quarks in the fermion loops, at values of the light-quark masses larger than the physical values and an extrapolation in these masses is required. The calculations are generally made at a fixed value of the lattice spacing and thus contain discretization errors. An estimate of the quenching errors is obtained by comparing quenched and unquenched results at similar values of the lattice spacing. These comparisons are complemented by theoretical estimates of this uncertainty obtained by using quenched and unquenched chiral perturbation theory techniques [43].

The question arising at this point is: what is the best model to describe the theoretical systematic errors in lattice calculations, and hence the assessment of the uncertainty? We refer to the general introduction of the inferential framework given in Section 3. As happens with systematic errors in experiments, this *completely* relies on the confidence of the experts about possible variations of an influence parameter, the effect of quenching or what would happen in passing from a perturbative order to the other or changing the renormalization scheme. These evaluations are unavoidably subjective, though *not arbitrary*, as long as we use the judgements of responsible experts for each input quantity. Using their judgements we commit ourselves too. Therefore, hereafter, when we state that a parameter lies in a certain range with uniform distribution, it means that, in practice, we are 100% confident that the parameter lies in that region, and that, for any choice of a sub-interval of half the width, we are in condition of indifference (i.e. 50% confidence) that the value of the parameter is inside the sub-interval or somewhere else. For a more extended discussion see for example [44] and references therein.

In conclusion, we cannot find any conceptual difference which would force us to treat experimental and theoretical uncertainties on a different footing and claim that the standard method is a perfectly justified scientific approach able to establish confidence levels for the quantities of interest. In the following, for each parameter, taken from lattice QCD evaluations, best estimates for its central value and attached uncertainties are given.

In order to check the stability of the results, we have also made the analysis with the flat part of the theoretical uncertainty increased by a factor two.

4.2 Evaluation of the parameter $f_{B_{\mathrm{d}}}\sqrt{\hat{B}_{B_{\mathrm{d}}}}$

Traditionally, the B-parameter of the renormalized operator is defined as

$$\langle \bar{B}_d | Q_d^{\Delta B=2}(\mu) | B_d \rangle = \frac{8}{3}\, m_{B_d}^2 f_{B_d}^2 \; B_{B_d}(\mu)\,, \tag{21}$$

where $Q_d^{\Delta B=2} = (\bar{b}\gamma^\mu(1-\gamma_5)d)(\bar{b}\gamma_\mu(1-\gamma_5)d)$ and μ is the renormalization scale. This definition stems from the vacuum saturation approximation (VSA) in which $B_{B_d} = 1$. Similarly one can define B_{B_s}. The renormalization group invariant B-parameter $\hat{B}_{B_d}$ of Equation (2) is defined as

$$\hat{B}_{B_d} = \alpha_s(\mu)^{-\gamma_0/2\beta_0} \left(1 + \frac{\alpha_s(\mu)}{4\pi} J\right) \; B_{B_d}(\mu)\ . \tag{22}$$

$\gamma_0 = 4$ in all schemes whereas J depends on the scheme used for renormalizing $Q_d^{\Delta B=2}(\mu)$. In the theoretical expressions, the physical amplitudes are always defined in terms of $\hat{B}_{B_d}$. The advantage is that this quantity is not only renormalization scale, but also renormalization-scheme independent. A residual scheme dependence remains only because the coefficient renormalizing the lattice operator is computed at a fixed order in perturbation theory (NLO in this case), whereas its matrix element is computed non-perturbatively. This problem would arise in any approach that computes physical amplitudes by combining the Wilson coefficients of the effective Hamiltonian, which are computed perturbatively, with hadronic matrix elements. Thus it is not specific to lattice calculations. The scheme dependence can be reduced by increasing the order at which Wilson coefficients are computed in *continuum* perturbation theory.

Indeed the important quantity is not the B-parameter itself but the combination $(f_{B_d}\sqrt{\hat{B}_{B_d}})^2$ which is used as an alias for the physical amplitude to which it is simply related by the factor $8m_{B_d}^2/3$. This is similar to the kaon B parameter. In that case, however, the decay constant is taken from experiments. For the B_d meson, instead, we have to rely on theory also for the decay constant. Since the calculation of f_{B_d} and $\hat{B}_{B_d}$

Quenched		f_{B_d} (MeV)	f_{B_s}/f_{B_d}	f_D (MeV)	f_{D_s} (MeV)
APE [45]	97	180(32)	1.14(8)	221(17)	237(16)
FNAL [46]	97	$164(^{+14}_{-11})(8)$	$1.13(^{+5}_{-4})$	$194(^{+14}_{-10})(10)$	$213(^{+14}_{-11})(11)$
JLQCD [47]	98	173(4)(12)	$\simeq 1.15$	197(2)(17)	224(2)(19)
MILC$^{(*)}$ [48]	98	$157(11)(^{+25}_{-9})(^{+23}_{-0})$	$1.11(2)(^{+4}_{-3})(3)$	$192(11)(^{+16}_{-8})(^{+15}_{-0})$	$210(9)(^{+25}_{-9})(^{+17}_{-1})$
APE [49]	99	$173(13)(^{+34}_{-2})$	1.14(2)(1)	$216(11)(^{+5}_{-4})$	$239(10)(^{+15}_{-0})$
APE [50]	00	$174(22)(^{+7}_{-0})(^{+4}_{-0})$	$1.17(4)(^{+0}_{-1})$	$207(11)(^{+3}_{-0})(^{+3}_{-0})$	$234(9)(^{+3}_{-0})(^{+2}_{-0})$
UKQCD$^{(**)}$ [51]	00	$218(5)(^{+5}_{-41})$	$1.11(1)(^{+5}_{-3})$	$220(3)(^{+2}_{-24})$	$241(2)(^{+7}_{-30})$
MILC [52]	00	173(6)(16)	1.16(1)(2)	$200(6)(^{+12}_{-11})$	$223(5)(^{+19}_{-17})$
CP-PACS [53]	00	188(3)(9)	$1.148(8)(46)(^{+39}_{-0})$	218(2)(15)	$250(1)(18)(^{+6}_{-0})$
Lellouch and Lin [54]	00	$177(17)(^{+22}_{-26})$	$1.15(2)(^{+3}_{-2})$	$210(10)(^{+20}_{-16})$	$236(8)(^{+20}_{-14})$
Unquenched		f_{B_d} (MeV)	f_{B_s}/f_{B_d}	f_D (MeV)	f_{D_s} (MeV)
MILC [52]	00	$191(6)(^{+24}_{-18})(^{+11}_{-0})$	1.16(1)(2)(2)	$215(5)(^{+17}_{-15})(^{+8}_{-0})$	$241(4)(^{+32}_{-31})(^{+9}_{-0})$
CP-PACS [53]	00	208(10)(11)	$1.203(29)(43)(^{+38}_{-0})$	225(14)(14)	$267(13)(17)(^{+10}_{-0})$

Table 2: Decay constants from recent lattice calculations. The errors are those of the original publications. Some of the numbers have been taken from the recent lattice reviews [38, 40]. $(*)$: By taking in the extrapolation the previous MILC data closer to the continuum limit only (corresponding to $\beta = 6/g_0^2 \geq 6.0$), one would find $f_{B_d} \sim 180$ MeV. This is, in our opinion, a better extrapolation of these data. $(**)$: This large value was found by using the Sommer parameter to calibrate the lattice spacing. Using the ρ mass, instead, they find $f_{B_d} = 186$ MeV. We do not understand the origin of this large difference.

are strongly correlated, the best way is to take the combination $f_{B_d}\sqrt{\hat{B}_{B_d}}$ from a single calculation rather then using f_{B_d} and $\hat{B}_{B_d}$ from different studies as often done in the literature.

The most recent calculations of the combination $f_{B_d}\sqrt{\hat{B}_{B_d}}$, in the quenching approximation, come from ref. [50], obtained with the non-perturbatively improved action and a non-perturbative renormalization of the lattice operators, and from ref. [54], with the mean-field improved action and perturbatively renormalized operators

$$\begin{aligned} f_{B_d}\sqrt{\hat{B}_{B_d}} &= (206 \pm 28 \pm 7)\,\mathrm{MeV} \quad [50]\,, \\ f_{B_d}\sqrt{\hat{B}_{B_d}} &= (211 \pm 21^{+27}_{-28})\,\mathrm{MeV} \quad [54]\,. \end{aligned} \tag{23}$$

These results correspond to the following values obtained in the same simulation

$$\begin{aligned} f_{B_d} &= (174 \pm 22^{+7+4}_{-0-0})\,\mathrm{MeV} \quad \hat{B}_{B_d} = 1.38 \pm 0.11^{+0.00}_{-0.09} \quad [50]\,, \\ f_{B_d} &= (177 \pm 17^{+22}_{-26})\,\mathrm{MeV} \quad \hat{B}_{B_d} = 1.41 \pm 0.06^{+0.05}_{-0.01} \quad [54]\,. \end{aligned} \tag{24}$$

The numbers above agree with (our) world averages of quenched determinations, based on the results given in Tables 2 and 3, which were used in [57] [5]

$$\begin{aligned} f_{B_d} &= (175 \pm 20)\,\mathrm{MeV} \\ \hat{B}_{B_d} &= 1.36 \pm 0.08 \pm 0.05\,, \end{aligned} \tag{25}$$

which can be combined to give

$$f_{B_d}\sqrt{\hat{B}_{B_d}} = (205 \pm 24 \pm 8)\,\mathrm{MeV}\,. \tag{26}$$

[5] The same quenched results for f_{B_d} is quoted in ref. [41].

LQCD Calculations		$B_{B_d}(m_b)$	$\hat{B}_{B_d}$	$\hat{B}_{B_s}/\hat{B}_{B_d}$
JLQCD [55]	96	0.85(6)	1.26(8)	$\simeq 1$
BBS [37]	98	0.96(12)	1.42(18)	$\simeq 1$
APE [50]	00	$0.93(8)(^{+0}_{-6})$	$1.38(11)(^{+0}_{-9})$	0.98(5)
Lellouch and Lin [54]	00	$0.95(4)(^{+3}_{-1})$	$1.41(6)(^{+5}_{-1})$	$0.99(2)(^{+0}_{-1})$
HQET		$B_{B_d}(m_b)$	$\hat{B}_{B_d}$	$\hat{B}_{B_s}/\hat{B}_{B_d}$
Gimenez and Reyes [56](APE data)	98	0.87(5)(3)(2)	1.29(8)(5)(3)	$--$
Gimenez and Reyes [56](UKQCD data)	98	0.85(4)(3)(2)	1.26(6)(5)(3)	$--$

Table 3: B-meson B parameters from recent lattice calculations. For comparison we have evolved results and errors of the original publications to a common scale.

Note that in this case we have used decay constants and B-parameters from different calculations.

Our average of f_{B_d} does not include the results obtained with NRQCD. The reason is that the quenched results from two different groups are incompatible between each other: ref. [58] found $f_{B_d} = 147(11)(^{+8}_{-12})(9)(6)$ MeV to be contrasted with the recent CP-PACS result [41] $f_{B_d} = 191(5)(11)$ MeV. Moreover [58] finds a rather large difference, of about 40 MeV (from 147 MeV $\to$ 186 MeV), between the quenched and unquenched case. This is not confirmed by the most recent results given in Table 2, which give differences of the order of 20 MeV. In our opinion, the situation with this approach is still rather confused and therefore we do not use the NRQCD results (until it will not be clarified). This applies also to the related calculations of the $\hat{B}_{B_{d,s}}$ parameters, which are computed within the same framework.

Our average for $\hat{B}_{B_d}$ has been computed by combining values obtained with heavy quark masses in the charm region, extrapolated to the B mesons (denoted as LQCD in Table 3), with those obtained in the HQET [56], at lowest order in the $1/m_b$ expansion. The systematic error has been estimated from the different values obtained by using only the LQCD results or by combining LQCD and HQET predictions.

The world average given above for f_{B_d} includes data obtained after extrapolation to the continuum, as well as results obtained with improved actions at small values of a (for which discretization errors are expected to be smaller). A completely non-perturbative determination of the axial current renormalization constant Z_A has been also performed in several cases, thus eliminating this source of errors.

In the case of the B parameter, no systematic continuum extrapolation has been attempted yet. In this case it is reassuring that results obtained with perturbative and non-perturbative renormalization techniques and for a variety of values of lattice spacing are so close. In the absence of any indication of large discretization errors, we ignore them in the following.

As far as quenching errors are concerned, there is a general agreement that the value of the decays constants increases in the unquenched case. This is supported by both theoretical estimates with quenched and unquenched chiral perturbation theory [43], and by explicit numerical calculations, see Table 2. The MILC [41] and CP-PACS [53] Collaborations find very consistent results,

$$
\begin{aligned}
f_{B_d}^{unq}/f_{B_d}^{quen} &= 1.12^{+0.16}_{-0.11}\,, \\
f_{B_d}^{unq}/f_{B_d}^{quen} &= 1.11 \pm 0.06\,,
\end{aligned}
\tag{27}
$$

respectively. Most unfortunately no unquenched determinations of the B-parameters, or even better of $f_B\sqrt{\hat{B}_B}$, have been presented yet. For the B-meson B parameters, chiral perturbation theory suggests that the unquenching error is at most of the order of 10%. This prediction should be supported by explicit numerical simulations which are missing at the moment.

By assuming 10% quenching uncertainty for $\hat{B}_{B_d}$, and using the results in (26) and (27), we arrive to

$$
f_{B_d}\sqrt{\hat{B}_{B_d}} = (230 \pm 25 \pm 20)\,\mathrm{MeV}. \tag{28}
$$

which has been used in our analysis. In our preliminary analysis (mostly based on quenched results) which has been presented in [57], we used instead

$$
f_{B_d}\sqrt{\hat{B}_{B_d}} = (220 \pm 25 \pm 20)\mathrm{MeV}\,, \tag{29}
$$

which is very close to the present value. The changes in the results for the relevant quantities ($\sin(2\alpha)$, $\sin(2\beta)$, γ, etc.), induced by the difference between (28) and (29), are negligible.

4.3 Evaluation of the parameter ξ

Besides the $B^0 - \bar{B}^0$ amplitude, an additional constraint is given by the ratio

$$
\frac{\Delta m_s}{\Delta m_d} = \frac{|V_{ts}|^2}{|V_{td}|^2}\left(\frac{m_{B_s}}{m_{B_d}}\right)\xi^2, \tag{30}
$$

where $\xi = f_{B_s}\sqrt{\hat{B}_{B_s}}/f_{B_d}\sqrt{\hat{B}_{B_d}}$. By combining the results for the decay constant ratios f_{B_s}/f_{B_d} given in Table 3 and for the B parameters (the latter being always very close to one) one finds always a number of ~ 1.15. As mentioned before, there is no confirmation, neither in the quenched nor in the unquenched case, of a value as large as $\xi = 1.3$ as found in [37]. With an estimate of the uncertainty on $\hat{B}_{B_s}/\hat{B}_{B_d}$ of 10% [43], we then find

$$
\xi = 1.14 \pm 0.03 \pm 0.05 \tag{31}
$$

which is the value used in our study. In our preliminary analysis we used, instead, $\xi = 1.14 \pm 0.06$ [57].

4.4 Evaluation of the parameter B_K

The kaon B-parameter, B_K, is one of the most studied, and more accurately known, quantities in lattice calculations. Very precise values have been obtained, within the quenching approximation [39]

$$
\begin{aligned}
&B_K^{\overline{\mathrm{MS}}}(\mu = 2\,\mathrm{GeV}) = 0.63 \pm 0.04 \quad [39] \\
&B_K^{\overline{\mathrm{MS}}}(\mu = 2\,\mathrm{GeV}) = 0.62 \pm 0.03 \quad [59]
\end{aligned}
\tag{32}
$$

which correspond to the renormalization group invariant B parameter

$$\hat{B}_K = 0.87 \pm 0.06\,. \tag{33}$$

This uncertainty, meant as standard deviation, includes the contribution from the statistical error and the deviation from the extrapolation to the continuum limit (this error is much larger than those on individual data at fixed lattice spacing). The physical amplitude has been computed at the NLO by using boosted perturbation theory [60]. A non-perturbative renormalization of the lattice operator is preferable but has not been performed yet. Previous experience on other quantities leads to an estimate of the error due to the use of the perturbative renormalization of the order of 5%.

Uncertainties, due to the quenching approximation, have been evaluated both theoretically and numerically. Using chiral perturbation theory, the error on B_K due to the quenched approximation has been estimated to be negligible for degenerate quark masses ($m_s = m_d$) and of the order of 5% for realistic quark masses [43]. A numerical unquenched calculation with $n_f = 2$ and $n_f = 4$ resulted in a shift upwards of the value of B_K by $(5 \pm 2)\%$ [61]. This calculation was performed at fixed lattice spacing and only the difference between the quenched and unquenched B_K for similar values of a was studied. Both the theoretical estimate and the numerical evaluation give a quenching error of $\mathcal{O}(5\%)$. The residual uncertainty is due to our ignorance on the dependence of this difference on the value of the lattice spacing. Since so far we have very accurate results with continuum extrapolation in the quenched case only and unquenched results without continuum extrapolation, it seems reasonable to consider a systematic uncertainty corresponding to a uniform distribution spanning the range of $\pm 15\%$ corresponding to the maximum discretization effect expected in the unquenched case [39]. For the central value, we have taken the quenched result given in Equation (33). We thus obtain

$$\hat{B}_K = 0.87 \pm 0.06 \pm 0.13 \tag{34}$$

which has been used in our study. This range of values is in good agreement with others which can be found in the literature [62, 63]. The allowed range for $\hat{B}_K$ in our evaluation and in [62, 63] is smaller than the one quoted in [21, 23], namely $0.6 \leq \hat{B}_K \leq 1.0$. That estimate tries to include results obtained with techniques different from lattice calculations, such as QCD Sum Rules and $1/N$ expansion. In our opinion, the other approaches lack of the accuracy and control of systematic effects reached by lattice calculations for this parameter. We think instead that Equation (34) corresponds to our best knowledge of $\hat{B}_K$, given the present understanding of the theory.

5 Other inputs

In this section we briefly discuss other inputs which have been used in the present analysis: $|V_{cb}|$, $|V_{ub}|$ and λ. These are obtained from experimental measurements combined with several theoretical predictions which are discussed below.

5.1 Extraction of $|V_{cb}|$

Using exclusive decays $\bar{B}^0_d \to D^{*+}\ell^-\bar{\nu}_\ell$, the value of $|V_{cb}|$ is obtained by measuring the differential decay rate at maximum q^2, which is the mass squared of the charged lepton-neutrino system. At $q^2 = q^2_{max}$, the D^{*+} is produced at rest in the $\bar{B}^0_d$ hadron rest frame and HQET can be invoked to obtain the value of the corresponding form factor $F_{D^*}(w=1)$. The variable w is usually introduced as the product of the four-velocities of the $\bar{B}_d^0$ and D^{*+} mesons

$$w = v_{\bar{B}_d^0} \cdot v_{D^{*+}} = \frac{m_{\bar{B}_d^0}^2 + m_{D^*}^2 - q^2}{2\, m_{\bar{B}_d^0}\, m_{D^*}}, \qquad w = 1 \text{ for } q^2 = q_{max}^2 \,. \tag{35}$$

In terms of w, the differential decay rate can be written as

$$\frac{dBR_{D^*}}{dw} = \frac{1}{\tau_{B_d^0}} \frac{G_F^2}{48\pi^3} m_{D^*}^3 (m_B - m_{D^*})^2 K(w) \sqrt{w^2-1} F_{D^*}^2(w)\, |V_{cb}|^2 \,, \tag{36}$$

where $K(w)$ is a kinematic factor.

As the decay rate is zero for $w = 1$, the w dependence has to be adjusted over the measured range.

Four measurements obtained by the LEP collaborations [64] have been averaged with the result from CLEO [65], taking into account correlations induced by common sources of systematic uncertainties. Before averaging, results and errors have been recalibrated using a common set of values for the external parameters, such as the $\bar{B}_d^0$ lifetime and charm hadron decay branching fractions [64].

The average value is

$$F(1)\ |V_{cb}| = (37.2 \pm 1.5) \times 10^{-3}, \quad \chi^2/NDF = 9.15/4 \,. \tag{37}$$

The "fit probability"[6] is only 6%, giving rise to the suspect that there could be some extra systematic effect playing an important role.

Another evaluation of this average has thus been done following a model [68] developed to combine results which appear to be in mutual disagreement. It consists in assuming that quoted uncertainties (s_i) for each measurement are proportional to the unknown real uncertainty ($\sigma_i/r_i = s_i$). A distribution probability for r_i is then assumed. A simple model, which depends on two parameters, δ and ζ is given by

$$f(r) = \frac{2\zeta^\delta r^{-(2\delta+1)} \exp^{-\zeta/r^2}}{\Gamma(\delta)} \,. \tag{38}$$

The natural choice for δ and ζ corresponds to an expected value of 1 for r_i, with 100% uncertainty. The final results are largely independent from the precise value of the parameters. We have chosen the values of $\zeta = 0.6$ and $\delta = 1.3$ taken from Ref. [68].

The probability distribution for the quantity $F(1)\ |V_{cb}|$ is then obtained using the Bayes theorem

$$f(x) \propto \int_0^\infty \exp\left[-\frac{1}{2}\sum_{i,j=1}^{5}(x-x_i) w_{i,j}(\sigma_i/r_i, \sigma_j/r_j)(x-x_j)\right] \left(\prod_{i=1}^{5} f(r_i)\, dr_i\right) \,. \tag{39}$$

In this expression, $x = F(1)\ |V_{cb}|$ and $w_{i,j}$ is the weight matrix. The latter is obtained by inverting the error matrix, after the inclusion of the uncertainties given by the quantities

[6] "p-value" would be the correct modern statistics term to be used instead of "fit probability" or "χ^2 probability" [66, 67]. These expressions can be highly misleading as explained in [32].

r_i (considered to be independent for the different measurements). From this distribution, which has non-Gaussian tails, the average and the standard deviation have been obtained

$$F(1)\ |V_{cb}| = (37.1 \pm 1.9) \times 10^{-3}. \tag{40}$$

The central value is practically the same as obtained with the usual fit and the standard deviation has increased by 30%.

From the exclusive measurements and using $F(1) = 0.88 \pm 0.05$ [69], $|V_{cb}|$ has been obtained

$$|V_{cb}| = (42.2 \pm 2.1 \pm 2.4) \times 10^{-3}\,. \tag{41}$$

The inclusive measurements of the semileptonic branching fraction of b-hadrons give instead

$$|V_{cb}| = (40.7 \pm 0.5 \pm 2.0) \times 10^{-3}\,. \tag{42}$$

The interested reader may consult [69] for more details on the quoted uncertainty for inclusive decays.

The combination of all results, using the procedure explained previously when averaging $F(1)\ |V_{cb}|$ measurements, taking into account correlated systematics, gives

$$|V_{cb}| = (41.0 \pm 1.6) \times 10^{-3}\,. \tag{43}$$

The corresponding p.d.f, which has been used in the present analysis, is shown in Figure 1.

5.2 Extraction of $|V_{ub}|$

$|V_{ub}|$ has been obtained from the measurements done at CLEO and LEP. The CLEO collaboration [70] has measured the branching fraction for the decay $\bar{B}^0_d \to \rho^+ \ell^- \bar{\nu}_\ell$ and deduced a value for $|V_{ub}|$ using several models to describe the decay form factors. LEP collaborations [71] have developed dedicated algorithms to be sensitive to a large fraction of the inclusive decay rate $b \to u \ell^- \bar{\nu}_\ell$ and, with some assumptions, a value for $|V_{ub}|$ is obtained. The two measurements are

$$\begin{aligned} |V_{ub}| &= (32.5 \pm 2.9 \pm 5.5) \times 10^{-4} \quad \text{CLEO} \\ |V_{ub}| &= (41.3 \pm 6.3 \pm 3.1) \times 10^{-4} \quad \text{LEP} \end{aligned} \tag{44}$$

where the second uncertainty is theoretical. The p.d.f. for the CLEO measurement is thus a convolution of a Gaussian and a flat distribution. The theoretical error for the LEP measurement, being the convolution of several different errors, is taken from a Gaussian distribution [71]. Combining the two distributions, in practice we obtain almost a Gaussian p.d.f. corresponding to

$$|V_{ub}| = (35.5 \pm 3.6) \times 10^{-4}. \tag{45}$$

5.3 The parameter λ

Measurements of $|V_{ud}|$ and $|V_{us}|$ reported in [29] have been combined using the procedure explained in Section 5.1, assuming that $|V_{ud}| = \cos\theta_c$ and $|V_{us}| = \sin\theta_c \equiv \lambda$. The additional contribution from $|V_{ub}|^2$ to the unitarity condition can be safely neglected owing